Solution 1: Hard Margin Classifier

(a) The function is symmetric, as we can prove in the following way:

$$k(\mathbf{x}, \tilde{\mathbf{x}}) \ = \mathbf{x}^T\tilde{\mathbf{x}} = \tilde{\mathbf{x}}^T\mathbf{x} = k(\tilde{\mathbf{x}}, \mathbf{x}) \tag{1}$$

To check if the kernel gram matrix is positive definite, we will remember the definition of $\mathbf{X}$

$$\mathbf{X} = \left\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\right\} \in \mathbb{R}^{p\times n} \tag{2}$$

Taking that into account, the kernel gram matrix is defined as :

$$\begin{aligned}\mathbf{K} &= \begin{pmatrix} \left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(1)}\right\rangle & \left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)}\right\rangle & \cdots & \left\langle \mathbf{x}^{(1)}, \mathbf{x}^{(n)}\right\rangle \\ \left\langle \mathbf{x}^{(2)}, \mathbf{x}^{(1)}\right\rangle & \left\langle \mathbf{x}^{(2)}, \mathbf{x}^{(2)}\right\rangle & \cdots & \left\langle \mathbf{x}^{(2)}, \mathbf{x}^{(n)}\right\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \left\langle \mathbf{x}^{(n)}, \mathbf{x}^{(1)}\right\rangle & \left\langle \mathbf{x}^{(n)}, \mathbf{x}^{(2)}\right\rangle & \cdots & \left\langle \mathbf{x}^{(n)}, \mathbf{x}^{(n)}\right\rangle \end{pmatrix} \\ &= \left\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\right\}^T \left\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}\right\} \\ &= \mathbf{X}^T\mathbf{X} \in \mathbb{R}^{n\times n}\end{aligned} \tag{3}$$

$$\begin{aligned}\mathbf{a}^T\mathbf{K}\mathbf{a} &= \mathbf{a}^T\mathbf{X}^T\mathbf{X}\mathbf{a} \\ &= (\underbrace{\mathbf{a}\mathbf{X}}_{\mathbf{z}})^T(\mathbf{a}\mathbf{X}) \\ &= \mathbf{z}^T\mathbf{z} \\ &= \|\mathbf{z}\|_2^2 > 0\end{aligned} \tag{4}$$

We can conclude that this function is a kernel.

(b) The function is not a kernel, as we can prove that the kernel gram matrix is not positive definite. We will use $x^{(1)} = \frac{\pi}{2}$ and $a = 1$. For this case, the kernel gram matrix is a scalar:

$$\begin{aligned}K = k(x^{(1)}, x^{(1)}) = \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = -1 \\ a\cdot K \cdot a = 1 \cdot -1 \cdot 1 = -1 < 0\end{aligned} \tag{5}$$

(c) The function is not a kernel, we can prove that the kernel gram matrix is not positive definite by using a counter-example, we will use $x^{(1)} = 1$ and $x^{(2)} = 2$, and $\mathbf{a} = (1, -1)^T$:

$$\mathbf{K} = \begin{pmatrix} max(x^{(1)}, x^{(1)}) & max(x^{(1)}, x^{(2)}) \\ max(x^{(2)}, x^{(1)}) & max(x^{(2)}, x^{(2)}) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix} \tag{6}$$

$$\begin{aligned}\mathbf{a}^T\mathbf{K}\mathbf{a} &= (1, -1) \cdot \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix} \cdot (1, -1)^T \\ &= (-1, 0) \cdot (1, -1)^T \\ &= -1\end{aligned} \tag{7}$$

(d) In this case, we can use properties of the kernels to prove that the function is indeed a kernel.

- Multiplying a kernel by a non-negative scalar always results in a kernel. Accordingly, we can say that $\alpha k_1(\mathbf{x}, \tilde{\mathbf{x}})$ and $\beta k_2(\mathbf{x}, \tilde{\mathbf{x}})$ are kernels.
- The sum of two kernels always results in a kernel. Accordingly, we can say that $\alpha k_1(\mathbf{x}, \tilde{\mathbf{x}}) + \beta k_2(\mathbf{x}, \tilde{\mathbf{x}})$ is a kernel.